PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12Q 1/68, C12N 5/10, C07K 14/00, 2/00, 4/00, 14/00, 16/00, C07H 21/04, A61K 39/395

A1

(11) International Publication Number: WO 96/22390

(43) International Publication Date: 25 July 1996 (25.07.96)

(21) International Application Number: PCT/US96/00058

(22) International Filing Date: 16 January 1996 (16.01.96)

(30) Priority Data:
08/374,445 17 January 1995 (17.01.95) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/374,445 (CIP)
Filed on 17 January 1995 (17.01.95)

(71) Applicant *(for all designated States except US)*: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES [US/US]; 10010 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: EVANS, Ronald, M. [US/US]; 1471 Cottontail Lane, La Jolla, CA 92037 (US). BLUMBERG, Bruce [US/US]; 3435 Lebon Drive #1122, San Diego, CA 92122 (US). UMESONO, Kazuhiko [JP/JP]; University Apartments, B402, 8916-5, Takayama, Ikoma, Nara 63-01 (JP).

(74) Agent: REITER, Stephen, E.; Pretty, Schroeder, Brueggemann & Clark, Suite 2000, 444 South Flower Street, Los Angeles, CA 90071 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHODS, POLYPEPTIDES, NUCLEOTIDE SEQUENCE OF XOR-6, A VITAMIN D-LIKE RECEPTOR FROM XENOPUS

(57) Abstract

The inventions disclosed are new members of the steroid receptor superfamily of receptors of which a representative member has been designated XOR-6. The receptors are responsive to hydroxy, mercapto or aminobenzoates and are expressed in Xenopus leavis embryos. XOR-6 is most closely related to the vitamin D3 receptor. The amino acid sequences are about 73 % identical in the DNA-binding domains and about 42 % identical in the ligand binding domain. Like the vitamin D3 receptor, XOR-6 has an extended D region between the DNA and ligand binding domains. The region amino-terminal to the XOR-6 DNA-binding domain is extremely acidic which may influence its ability to activate genes. Southern blots show that XOR-6 related sequences are present in other vertebrates including humans. Also disclosed are nucleotide sequences encoding the XOR-6 receptor, constructs and cells containing sane, and probes derived from the XOR-6 sequence. Hydroxy, mercapto and aminobenzoates modulate the transcription of the invention receptors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

METHODS, POLYPEPTIDES, NUCLEOTIDE SEQUENCE OF XOR-6, A VITAMIN D-LIKE RECEPTOR FROM XENOPUS

FIELD OF THE INVENTION

The present invention relates to intracellular receptors, and ligands therefor. In a particular aspect, the present invention relates to methods for the modulation of processes mediated by invention receptors, as well as methods for the identification of compounds which effect such modulation.

BACKGROUND OF THE INVENTION

Nuclear receptors constitute a large superfamily of ligand-activated transcription factors. Members of this family influence transcription either directly, through specific binding to the promoters of target genes (see Evans, in *Science* 240:889-895 (1988), or indirectly, via protein-protein interactions with other transcription factors (see, for example, Jonat et al., in *Cell* 62:1189-1204 (1990), Schuele et al., in *Cell* 62:1217-1226 (1990), and Yang-Yen et al., in *Cell* 62:1205-1215 (1990)). The steroid/thyroid receptor superfamily includes receptors for a variety of hydrophobic ligands including cortisol, aldosterone, estrogen, progesterone, testosterone, vitamin D_3, thyroid hormone and retinoic acid, as well as a number of receptor-like molecules, termed "orphan receptors" for which the ligands remain unknown (see Evans, 1988, supra). These receptors all share a common structure indicative of divergence from an ancestral archetype.

Identification of ligands for orphan receptors presents a significant challenge for the future since the number of orphan receptors which have been identified far exceeds the number of receptors with known ligands. Indeed, at least 40 genes, both vertebrate and invertebrate, have been identified which are structurally related to the steroid/thyroid receptor superfamily, but whose ligands are unidentified. Among these are *Drosophila* genes of known developmental significance including: the gap gene, *knirps* (Nauber et al., in *Nature* 336:489-492 (1988), the terminal gene *tailless*, involved in patterning the head and tail regions (Pignoni et al., in *Cell* 62:151-163 (1990), *seven-up*, which influences photoreceptor cell-fate (Mlodzik et al., in *Cell* 60: 211-224 (1990), and *ultraspiracle*, a gene required both maternally and zygotically for pattern formation (Oro et al., in *Nature* 347: 298-301 (1990)).

The identification of important *Drosophila* developmental genes as members of the steroid/thyroid hormone receptor superfamily suggests that vertebrate orphan receptors will have important developmental functions. Furthermore, the identification of ligands for orphan receptors could lead to the discovery of novel morphogens, teratogens and physiologically important hormones.

BRIEF DESCRIPTION OF THE INVENTION

In accordance with the present invention, we have identified new members of the steroid receptor superfamily of receptors, a representative member of which has been designated XOR-6. Invention receptors are responsive to hydroxy, mercapto or amino benzoates, and are expressed, for example, in *Xenopus laevis* embryos. XOR-6 is most closely, although distantly, related to the vitamin D3 receptor (VDR). The proteins are about 73% identical in

amino acid sequence in the DNA-binding domains and about 42% identical in the ligand binding domain. Like VDR, XOR-6 has an extended D region between the DNA and ligand binding domains. Notably, the region amino-terminal to the XOR-6 DNA-binding domain is extremely acidic. This may influence its ability to activate target genes. XOR-6 is not restricted to *Xenopus* because southern blots show the presence of XOR-6-related sequences in a variety of other vertebrates. Indeed, a human genomic clone for an XOR-6 related gene has recently been isolated.

In accordance with a particular aspect of the present invention, there are also provided nucleic acid sequences encoding the above-identified receptors, as well as constructs and cells containing same, and probes derived therefrom. Furthermore, we have also discovered that hydroxy, mercapto or amino benzoates modulate the transcription activating effects of invention receptors.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 presents a schematic comparison between XOR-6 and the human vitamin D3 receptor. The two amino acid sequences were aligned using the program GAP (see Devereaux et al., in *Nucl. Acids Res.* 12:387-395 (1984)). Similarity between XOR-6 and hVDR is expressed as percent amino acid identity.

Figure 2 demonstrates that XOR-6 and hRXRα interact *in vivo*. The plasmids indicated in the figure were co-transfected into CV-1 cells along with the reporter tk(galp)3-luc and CMX-βgal. Note the strong suppression of basal transcription when GAL-XOR6 was added (right panel). This is characteristic of previously characterized ligand-dependent RXR heterodimeric partners.

Figure 3 illustrates the activation of XOR-6 by a variety of amino benzoate derivatives. Thus, 10^{-6}M of each compound was tested in the co-transfection assay for its ability to activate GAL-XOR6. Comparable results were obtained with full-length XOR-6.

Figure 4 illustrates the interaction of XOR-6 and RA signalling pathways, specifically demonstrating the synergism between partially purified XOR-6 agonist and the RXR ligand 9-cis RA. Receptors were transfected into cells and incubated with the indicated concentrations of agonists.

Figure 5 illustrates the interaction of XOR-6 and RA signalling pathways, specifically demonstrating how the Overexpression of full-length XOR-6, or the GAL-XOR-6 construct, interferes with retinoic acid (RA) signalling through the RARβ-RARE. 1 μg of XOR-6 expression plasmid was co-transfected into CV-1 cells with 5 μg of tk-β REx2-luc, and challenged with the indicated concentrations of all-trans retinoic acid.

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, we have identified new members of the steroid receptor superfamily of receptors, a representative member of which has been designated XOR-6. Invention receptors are responsive to hydroxy, mercapto or amino benzoates, and are expressed, for example, in *Xenopus laevis* embryos. Invention receptor comprises a protein of approximately 386 amino acids (see SEQ ID NO:2), which is most closely, although distantly, related to the vitamin D3 receptor (VDR). Also provided herein is a 2191 bp cDNA which encodes an example of invention receptors (see SEQ ID NO:1).

XOR-6 and VDR are about 73% identical in amino acid sequence in the DNA-binding domains and about 42% identical in the ligand binding domain. Like the VDR, XOR-6 has an extended D region between the DNA and ligand binding domains. Notably, the region amino-terminal to the XOR-6 DNA-binding domain is extremely acidic. This may influence its ability to activate target genes. XOR-6 is not restricted to *Xenopus* because southern blots show the presence of XOR-6-related sequences in a variety of other vertebrates.

XOR-6 was discovered as part of a search for nuclear receptors expressed early in *Xenopus laevis* development. Thus, cDNAs encoding transcripts from nine different genes were isolated. These included xRARα, xRARγ, xRXRα, xRXRγ and five different orphan receptors. The presence of this diversity of receptors early in development suggests that their ligands might play important roles in morphogenetic signalling processes. Therefore it was of particular interest to identify those orphan receptors which had a high probability of showing ligand dependence.

Because most known RXR heterodimeric partners are ligand responsive, the above-described orphan receptor collection was screened for the ability to heterodimerize with RXR both *in vitro* and *in vivo*. One such orphan receptor, XOR-6 (for *Xenopus* Orphan Receptor 6). XOR-6 is a novel heterodimeric partner for RXR both *in vitro* and *in vivo*, further extending the family of nuclear receptors which require RXR for high-efficiency DNA-binding. XOR-6:RXR heterodimers apparently prefer to bind direct repeats separated by four nucleotides (DR-4), as does the thyroid hormone receptor. XOR-6 expression significantly blunts the ability of RAR to activate gene expression suggesting that these two signalling pathways block each

other's ability to activate gene expression perhaps by influencing their common heterodimeric partner, RXR.

Based on the presumption that XOR-6 and its ligand must be co-expressed at some time during development, an unbiased, bioassay directed screen for XOR-6 agonists in HPLC fractionated organic extracts derived from a mixture of developmental stages was undertaken. A potent agonist was purified, and identified as 3-amino-ethyl-benzoate (3-AEB). Specific binding of 3-AEB to XOR-6 has been demonstrated herein, identifying it as a true ligand for this receptor. Additional ligands for XOR-6, e.g., hydroxy benzoates and mercapto benzoates, have also been identified. Accordingly, XOR-6 and ligands therefor represent a hitherto unknown hormonal signalling pathway.

RNAse protection assays were employed to measure steady-state mRNA levels over a developmental time sequence. XOR-6 mRNA is present in the unfertilized egg and remains at a relatively constant level until after gastrulation. It persists thereafter at a much reduced level until at least stage 45. To investigate whether XOR-6 mRNA is localized in the pre-midblastula embryo, blastulae were dissected into three major components, the animal cap, marginal zone and endoderm. RNAse protection analysis showed that there is no obvious localization of the maternally encoded XOR-6 mRNA at this stage.

Zygotic transcripts first become noticeable during neurulation (stage 14) where they appear in the anterior neural folds and the region lateral thereto. As the neural folds close, staining becomes more medial until finally appearing as an inverted Y at about stage 20. This is exactly the same pattern as cells which give rise to the hatching gland. Interestingly, this staining pattern defines boundaries of the future head. By stage 38, XOR-6

mRNA is restricted to the head, but is not limited to the hatching gland.

In vitro DNA-binding studies were used to determine the DNA-binding specificity of XOR-6. XOR-6 and hRXRα are seen to heterodimerize and bind DNA in a cocktail of response elements. This binding is strongly cooperative, as neither receptor alone showed DNA-binding at the protein concentrations used in the assay. This binding is also specific to hRXRα, because hRARα does not enhance XOR-6 DNA binding. Similar results are obtained using xRXRα.

A finer analysis of XOR-6:hRXRα binding specificity shows that the heterodimer binds to a subset of the known response elements in the cocktail: it binds weakly to DR-3 (but not the osteopontin vitamin D response element (SPP-VDRE), which is a variant of DR-3), strongly to DR-4 (and the murine leukemia virus (MLV-TRE), a DR-4-like element), and weakly to DR-5 (but strongly to the RARβ response element, a DR-5-like element). No significant binding is seen to synthetic or natural response elements corresponding to DR-0,1,2 or 6 (i.e., direct repeats having spacers of 0, 1, 2 or 6 nucleotides, respectively). These data indicate that the XOR-6:hRXRα heterodimer prefers to bind a DNA sequence consisting of directly-repeated AGTTCA half sites, separated by four nucleotides.

It was next tested to determine whether the XOR6:xRXRα heterodimer exhibited the predicted DNA-binding specificity. *In vitro* transcribed, translated XOR-6 and xRXRα proteins were tested for binding to direct repeats of AGTTCA separated by 1, 2, 3, 4, or 5 nucleotides (see Perlmann et al., in *Genes Dev.* 7:1411-1422 (1993)). The heterodimer is observed to exhibit the expected binding specificity to a response element comprising two half-sites (each having the sequence AGTTCA) separated by 4

nucleotides. This allowed the design of a specific XOR-6 reporter gene, tk-X6RE-luc (wherein the response element has the sequence AGTTCA TGAG AGTTCA; SEQ ID NO:3), which can be activated by XOR-6 in the presence of HPLC-purified embryo extracts.

In order to demonstrate that XOR-6 and RXR interact in vivo, a modification of the two hybrid system (see Fields and Song, in *Nature* 340:245-246 (1989), or Nagpal et al., in *Cell* 70:1007-1019 (1992)) was employed. This system relies on functional dimeric interactions between two proteins, one carrying the ability to bind a particular DNA-response element, and the other carrying the transactivation function, to reconstitute DNA-binding and transcriptional activation in a single complex.

Applying this system to XOR-6 and RXR, VP16-hRXRα (a constitutive activator), GAL-XOR-6 and $tk(gal_p)_3$-luc were employed. Functional interaction between XOR-6 and hRXRα should lead to constitutive activation of the reporter gene when all three constructs are transfected together. VP16-hRXRα alone does not activate the reporter because it lacks the ability to bind to a GAL4 response element. Activation of the reporter occurs only when GAL-XOR-6 and VP16-hRXRα are cotransfected. Moreover, GAL-XOR-6 shows strong suppression of reporter gene basal activity (see Figure 2), which parallels effects elicited by GAL-hRARα, GAL-hTRβ and GAL-hVDR. Based on these observations, it can be concluded that XOR-6 and hRXRα can form functional heterodimers *in vivo*, that GAL-XOR-6 is unable to activate target genes in the absence of its ligand, and that unliganded GAL-XOR6, like most other ligand-dependent RXR partners, suppresses basal activity of a reporter construct to which it can bind.

To demonstrate that XOR-6 hormone responsiveness differs from that of other RXR dimeric partners (e.g., RAR,

VDR, TR, and PPAR), the response of GAL-XOR-6 to agonists for the above receptors was tested. GAL-XOR-6 was not activated by a cocktail containing thyroid hormone (10^{-7}M), vitamin D3 (10^{-7}M), all-trans RA (10^{-6}M), or the peroxisome proliferator WY-14,643 ($5x10^{-6}$M), while GAL-VDR, GAL-hRARα, GAL-hTRβ, and GAL-mPPARα are activated by the cocktail. It can be concluded, therefore, that XOR-6 defines a novel RXR-dependent, ligand-mediated signalling pathway.

A search for the XOR-6 ligand was instituted based on the presumption that the receptor and its ligand must be co-expressed at some time during development. Accordingly, an unbiased, bioassay directed screen for XOR-6 agonists was undertaken in HPLC fractionated organic extracts derived from a mixture of developmental stages. Total lipid extracts from a mixture of embryonic stages from fertilized eggs through swimming tadpoles were prepared and tested for the ability to activate both GAL-XOR6 or full-length XOR-6 in transfected CV-1 cells.

The total extract was partitioned between iso-octane and MeOH and again tested for bioactivity. Since the methanol phase contained most of the activity, it was further partitioned between ethyl acetate and H_2O. The ethyl acetate phase was shown to contain most of the activity and was thus further purified by reverse phase HPLC using several solvent systems. Absorbance was monitored between 200 and 600 nm, fractions were collected, dried and tested in the cotransfection assay (see, for example, U.S. Patent No. 5,071,773) for their ability to activate full-length and GAL-XOR6. The eluted, purified agonist was subjected to high resolution mass spectroscopy which yielded a mass/charge ratio of 165.19 daltons. This predicted a molecular formula of $C_9H_{11}O_2N$, which most closely matches the ethyl ester of amino benzoic acid (AEB). The fragmentation pattern in Electron Impact mass spectroscopy suggests the meta isomer of AEB as the predominant form.

The ortho, meta and para amino ethyl benzoates were tested for agonist activity in the cotransfection assay. All three activated XOR-6 with a rank order potency as follows:

3-AEB > 4-AEB >> 2-AEB.

3-AEB co-chromatographed with purified agonist and gave an identical UV spectrum to authentic 3-AEB. Thus, 3-AEB is unequivocally identified as the purified agonist. Moreover, 3-AEB specifically activates XOR-6 alone among an extensive collection of published and unpublished vertebrate nuclear receptors.

In order to investigate ligand binding, the protease protection assay described by Leng et al., in *J. Ster. Bioch. and Mol. Biol.* 46:643-661 (1993) and Keidel et al, in *Mol. Cell. Biol.* 14:287-298 (1994) was utilized. Thus, ^{35}S-labelled *in vitro* transcribed translated protein was incubated with increasing concentrations of various proteases in the presence of solvent carrier or the putative ligand. The presence of 3-AEB results in some protection from trypsin cleavage with a concomitant increase in the intensity of the intermediately sized cleavage products. This result is not seen in parallel experiments with xRARα or xRXRα, again suggesting specificity in ligand binding.

It was next attempted to determine whether compounds related to 3-AEB might also function as ligand for invention receptor. One likely candidate is the vitamin, 4-amino-benzoic acid (PABA). It was not possible, however, to demonstrate XOR-6 activation by 2-, 3-, or 4-amino benzoic acids, or the related 2-, 3-, or 4-amino salicylic acids. It is possible that the cell membrane is much less permeable to the acids than to the more lipophilic esters. This possibility was tested by comparing the activation by a series of esters differing in the length of the alkyl group. As shown in Fig 3, the more

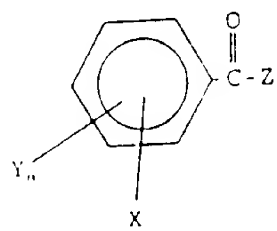
lipophilic esters showed increased activation with a rank order potency of 4-amino-butyl benzoate > 3-amino-ethyl benzoate > 4-amino-ethyl benzoate >> 4-amino methyl benzoate. These results suggest that the limiting step in XOR-6 activation is the transport of the ligand through the cell membrane. In conjunction with these studies, additional substituted benzoates, e.g., hydroxy benzoates and mercapto benzoates, have also been identified as ligands for invention receptor.

A potentially significant property of the XOR6:xRXRα heterodimer is its responsiveness to two ligands. Thus, in co-transfection experiments, either 9-cis RA or the partially purified agonist stimulated reporter gene expression in a receptor dependent manner. Unlike the response of RAR, VDR and TR heterodimers with RXR, which show additive effects on transcription, the XOR-6 ligand synergizes with 9-cis retinoic acid to activate its reporter gene (see Figure 4), reminiscent of the situation with PPAR (see Kliewer et al., in *Nature* 358:771-774 (1992)). This synergism occurs at several dilutions of the XOR-6 agonist and concentrations of 9-cis RA (see Figure 4). The demonstration of another heterodimer with dual hormone-responsiveness suggests that nuclear receptor heterodimers can generate combinatorial diversity by creating complexes with both novel DNA-binding properties and multiple hormonal activation levels. Such complexes would be ideal candidates for responding to combinations of graded morphogenetic signals during development.

Because XOR-6:RXR heterodimers bind well to a retinoic acid response element, βRARE, it was tested whether overexpression of XOR-6 could influence retinoic acid signalling through this element. As shown in Figure 5, it is found that co-expression of XOR-6 and βRARE significantly blunts the retinoic acid-responsiveness of

this promoter in a dose-dependent manner. This effect was strongest with full-length XOR-6 (24% of wild-type activity) but still detectable with GAL-XOR-6 (44% of wild-type activity). This suggests that maximal repression results from binding of XOR-6:RXR heterodimers to the βRARE, producing a non-productive transcription complex. The weaker inhibition by GAL-XOR-6 (which cannot bind to βRARE) suggests that sequestration of RXR in heterodimers unresponsive to retinoic acid also plays an inhibitory role.

In accordance with another embodiment of the present invention, there are provided a class of hydroxy, mercapto or amino benzoate compounds which are capable of acting as ligands for invention receptors. As employed herein, the phrase "hydroxy, mercapto or amino benzoate(s)" embraces compounds having the structure:



wherein

X is an hydroxy, alkoxy (of a lower alkyl, i.e., having 1-4 carbon atoms), mercapto, thioalkyl (of a lower alkyl), amino, alkylamino or acylamino group at the 2-, 3-, or 4-position of the ring,

each Y, when present, is independently selected from hydroxy, alkoxy, mercapto, thioalkyl, halide, trifluoromethyl, cyano, nitro, amino, carboxyl, carbamate, sulfonyl, sulfonamide, and the like,

Z is selected from -OR' or -NHR', wherein R' is selected from hydrogen, C_1-C_{12} alkyl, or C_5-C_{10} aryl, and

n is 0-2.

Presently preferred compounds embraced by the above generic formula include those wherein X is 2-, 3-, or 4-hydroxy or 3- or 4-amino, Z is alkoxy (i.e., methoxy, ethoxy or butoxy) and n is 0.

In accordance with yet another embodiment of the present invention, there are provided nucleic acids which encode the above-described receptor polypeptides. Exemplary DNAs include those which encode substantially the same amino acid sequence as shown in SEQ ID NO:2 (e.g., a contiguous nucleotide sequence which is substantially the same as nucleotides 166 - 1324 shown in SEQ ID NO:1). Preferred DNAs include those which encode the same amino acid sequence as shown in SEQ ID NO:2 (e.g., a contiguous nucleotide sequence which is the same as nucleotides 166 - 1324 shown in SEQ ID NO:1).

As used herein, nucleotide sequences which are substantially the same share at least about 90% identity, and amino acid sequences which are substantially the same typically share more than 95% amino acid identity. It is recognized, however, that proteins (and DNA or mRNA encoding such proteins) containing less than the above-described level of homology arising as splice variants or that are modified by conservative amino acid substitutions (or substitution of degenerate codons) are contemplated to be within the scope of the present invention.

In accordance with still another embodiment of the present invention, there are provided DNA constructs comprising the above-described DNA, operatively linked to regulatory element(s) operative for transcription of said

DNA and expression of said polypeptide in an animal cell in culture. There are also provided cells containing such construct, optionally containing a reporter vector comprising:

(a) a promoter that is operable in said cell,

(b) a hormone response element, and

(c) DNA encoding a reporter protein,

wherein said reporter protein-encoding DNA is operatively linked to said promoter for transcription of said DNA, and

wherein said promoter is operatively linked to said hormone response element for activation thereof.

In accordance with a still further embodiment of the present invention, there are provided probes comprising labeled single-stranded nucleic acid, comprising at least 20 contiguous bases in length having substantially the same sequence as any 20 or more contiguous bases selected from bases 1 - 2150, inclusive, of the DNA illustrated in SEQ ID NO:1, or the complement thereof. An especially preferred probe of the invention comprises at least 20 contiguous bases in length having substantially the same sequence as any 20 or more contiguous bases selected from bases 473 - 1324, inclusive, of the DNA illustrated in SEQ ID NO:1, or the complement thereof.

Those of skill in the art recognize that probes as described herein can be labelled with a variety of labels, such as for example, radioactive labels, enzymatically active labels, fluorescent labels, and the like. A presently preferred means to label such probes is with ^{32}P. Such probes are useful, for example, for the identification of receptor polypeptide(s) characterized by being responsive to the presence of hydroxy, mercapto or amino benzoate(s) to regulate the transcription of associated gene(s), said method comprising hybridizing test

DNA with a probe as described herein under high stringency conditions (e.g., contacting probe and test DNA at 65°C in 0.5 M $NaPO_4$, pH 7.3, 7% sodium dodecyl sulfate (SDS) and 5% dextran sulfate for 12-24 hours; washing is then carried out at 60°C in 0.1xSSC, 0.1% SDS for three thirty minute periods, utilizing fresh buffer at the beginning of each wash), and thereafter selecting those sequences which hybridize to said probe.

In another aspect of the invention, the above-described probes can be used to assess the tissue sensitivity of an individual to hydroxy, mercapto or amino benzoates by determining XOR-6 mRNA levels in a given tissue sample. It is expected that an individual having a high level of XOR-6 mRNA (or protein) will be sensitive to the presence of significant levels of amino benzoates, such as are used in sunscreen applications.

In accordance with yet another embodiment of the present invention, there are provided antibodies which specifically bind the above-described receptor polypeptides. Preferably, such antibodies will be monoclonal antibodies. Those of skill in the art can readily prepare such antibodies having access to the sequence information provided herein regarding invention receptors.

Thus, the above-described antibodies can be prepared employing standard techniques, as are well known to those of skill in the art, using the invention receptor proteins or portions thereof as antigens for antibody production. Both anti-peptide and anti-fusion protein antibodies can be used (see, for example, Bahouth et al. *Trends Pharmacol Sci*. 12:338-343 (1991); Current Protocols in Molecular Biology (Ausubel et al., eds.) John Wiley and Sons, New York (1989)). Factors to consider in selecting portions of the invention receptors for use as immunogen

(as either a synthetic peptide or a recombinantly produced bacterial fusion protein) include antigenicity, uniqueness to the particular subtype, and the like.

The availability of such antibodies makes possible the application of the technique of immunohistochemistry to monitor the distribution and expression density of invention receptors. Such antibodies could also be employed for diagnostic and therapeutic applications.

In accordance with yet another embodiment of the present invention, there is provided a method of testing a compound for its ability to regulate transcription-activating effects of invention receptor polypeptide(s), said method comprising assaying for the presence or absence of reporter protein upon contacting of cells containing said receptor polypeptide and reporter vector with said compound;

wherein said reporter vector comprises:

(a) a promoter that is operable in said cell,

(b) a hormone response element, and

(c) DNA encoding a reporter protein,

wherein said reporter protein-encoding DNA is operatively linked to said promoter for transcription of said DNA, and

wherein said promoter is operatively linked to said hormone response element for activation thereof.

Hormone response elements suitable for use in the above-described assay method comprise two half sites (each having the sequence AGTTCA), separated by a spacer of 3, 4 or 5 nucleotides. Those of skill in the art recognize that any combination of 3, 4 or 5 nucleotides can be used as the

spacer. Response elements having a spacer of 4 nucleotides (e.g., SEQ ID NO:3) are presently preferred.

Optionally, the above-described method of testing can be carried out in the further presence of ligand for invention receptors (e.g., a hydroxy, mercapto or amino benzoate), thereby allowing the identification of antagonists of invention receptors. Those of skill in the art can readily carry out antagonist screens using methods well known in the art. Typically, antagonist screens are carried out using a constant amount of agonist, and increasing amounts of a putative antagonist.

In accordance with a still further embodiment of the present invention, there is provided a method for modulating process(es) mediated by invention receptor polypeptides, said method comprising conducting said process(es) in the presence of at least one hydroxy, mercapto or amino benzoate (as defined hereinabove).

As shown herein, XOR-6 and RXR functionally interact both *in vitro* to preferentially bind a DR-4 type response element, and *in vivo* to activate a GAL4-based reporter in the two-hybrid assay. Thus a functional interaction has been identified between RXR and an orphan receptor within the cell to activate a reporter gene. This observation can be exploited to develop a high-sensitivity assay system for the XOR-6 ligand and for orphan receptor ligands in general, at least for those which interact with RXR.

The invention will now be described in greater detail by reference to the following non-limiting examples.

Example 1
cDNA isolation and characterization

XOR-6 was identified in a screen for maternally-expressed nuclear hormone receptors (Blumberg et al., in *Proc. Natl. Acad. Sci. USA* 89:2321-2325 (1992). Three clones were identified from an egg cDNA library, an additional two were isolated from a dorsal blastopore lip cDNA library. The longest clone was sequenced completely on both strands using a combination of directed subcloning and specific oligonucleotide priming. DNA sequences were compiled and aligned using the programs of Staden (Staden, in *Nucleic Acids Res.* 14:217-231 (1986), University of Wisconsin Genetics Computer Group (Devereaux et al., 1984, supra, and Feng and Doolittle (Feng and Doolittle, in *J. Mol. Evol.* 25:351-360 (1987). Database searching was performed using the BLAST network server at the National Center for Biotechnology Information (Altschul et al., *J. Mol. Biol.* 215:403-410 (1990)).

Example 2
RNA preparation and analysis

RNA was prepared from fertilized *Xenopus laevis* eggs and staged embryos as described by Blumberg et al., 1992, supra. The temporal and spatial patterns of expression were determined using RNAse protection as described by Blumberg et al., 1992, supra. The RNAse protection probes used are the following: EF-1α, nucleotides 790-1167; XOR-6, nucleotides 1314 to 1560, which represents the last three amino acids of the protein and part of the 3' untranslated region.

RNAse protection was performed with total RNA from the total ovary (10 μg); unfertilized egg (40 μg); 2-cell (40 μg); blastula (40 μg); gastrula (st 10, 10 μg), st 11, 8 μg); neurula (4 μg); tailbud (4 μg); swimming

tadpole (4 μg). Alternatively, RNAse protection was performed with 20 μg of total RNA from whole embryos or dissected animal caps, marginal zone, and vegetal pole.

A lateral view of a stage 12 embryo hybridized with antisense XOR-6 reveals that hybridization extends from the anterior-most end of the involuting mesoderm to the dorsal blastopore lip.

For localization studies, stage 8-9 embryos were dissected into animal, marginal and vegetal fragments and RNA was prepared using a proteinase K method as described by Cho et al., in *Cell* 65:55-64 (1991). Whole-mount *in situ* hybridization was performed as described by Harland, (1991). The entire cDNA shown in SEQ ID NO:1 was used as a probe for *in situ* hybridization. To make anti-sense RNA, the Bluescript II SK-plasmid containing the cDNA was linearized with SmaI and transcribed with T7 RNA polymerase. To produce sense RNA, the plasmid was digested with EcoRV and transcribed with T3 RNA polymerase.

Example 3

In vitro DNA-binding

DNA-binding analysis was performed using *in vitro* transcribed, translated proteins (Perlmann et al., 1993, supra. Oligonucleotides employed have been described previously (see Umesono et al., in *Cell* 65:1255-1266 (1990) and Perlmann et al., 1993, supra).

Thus, *in vitro* transcribed and translated proteins were mixed with a cocktail of hormone response elements containing DR0, DR1, PPRE, DR2, MLV-TRE, SPP1, and β-RARE. Thus, XOR-6 and hRXRα proteins were mixed and incubated with radiolabelled response elements. DR-1 through 5 are direct repeats of the sequence AGTTCA separated by 1-5 nucleotides. Reaction conditions and gel

electrophoresis employed were as described by Perlmann et al., 1993, supra.

Example 4
Cell culture and transfection studies

A suitable eukaryotic expression vector for use herein was constructed from the commercially available vector pCDNAI-AMP (Invitrogen). This vector allows expression from the strong cytomegalovirus early promoter, and bacteriophage T7 and SP6 promoter-driven production of sense and antisense RNA, respectively.

The cloning strategy employed was as follows: the three endogenous NcoI sites were removed by site directed mutagenesis, the polylinker region between XhoI and XbaI was removed by double digestion, endfilling and self ligation. A cassette consisting of the *Xenopus* β-globin leader and trailer derived from the plasmid pSP36T (see Amaya et al., in *Cell* 66:257-270 (1991)), separated by a synthetic polylinker (containing unique sites for NcoI, SphI, EcoRI, SalI, EcoRV, BamHI, and XbaI) was inserted between HindIII and NotI sites in the vector. The resulting plasmid, designated pCDG1, can be linearized with NotI to produce mRNA from the bacteriophage T7 promoter. The XOR-6 protein coding region was cloned between the NcoI and BamHI sites of pCDG1 and designated pCDG-XOR6.

pCMX-GAL4-XOR6 was constructed by cloning nucleotides encoding amino acids 103 to 386 of XOR-6 into the SalI to XbaI sites of pCMX-GAL4 (see USSN 08/177,740).

pCMX-VP16 receptor chimeras were constructed by fusing the potent VP16 transactivation domain (see Sadowski et al., in *Nature* 335:563-564 (1988)) to the amino terminus of the full-length hRXRα (see Mangelsdorf et al., *Nature* 345:224-229 (1990)), hRARα (see Giguere et al., in *Nature*

330:624-629 (1987)), or VDR (see McDonnell et al., in *Mol. Endocrinol.* 3:635-644 (1989)) protein coding regions.

CV-1 cells were maintained in DMEM containing 10% resin-charcoal stripped fetal bovine serum. Liposome-mediated transient transfections were performed using DOTAP reagent (Boehringer Manheim) at a concentration of 5 μg/ml in Opti-MEM (Gibco). After 12-18 hours, the cells were washed and fresh DMEM-10% serum was added, including receptor agonists if required. After a further 48 hour incubation, the cells were lysed and luciferase reporter gene assays and β-galactosidase transfection control assays performed. Reporter gene expression is normalized to the β-galactosidase transfection control and expressed as relative light units per O.D. per minute of β-galactosidase activity.

Example 5
Organic Extraction and HPLC analysis

Fresh or flash frozen embryos were homogenized in a large volume of 50% CH_2Cl_2/50% MeOH, typically 10 ml/gram of tissue. Denatured proteins were removed by filtration through diatomaceous earth and the liquid phase recovered and evaporated to dryness with a Buchi rotary evaporator. The resulting material was resuspended in a minimum volume of iso-octane and transferred to a separatory funnel. Non-polar and polar compounds were separated by partitioning between large volumes of iso-octane and MeOH. An agonist of XOR-6 partitioned primarily into the methanol layer.

The methanol phase was then dried, weighed, and partitioned between ethyl acetate and H_2O. An agonist for XOR-6 partitioned greater than 95% into ethyl acetate. The ethyl acetate phase was then dried, weighed, and fractionated by reverse phase HPLC, using several solvent systems.

Initially, the ethyl acetate phase was separated by isocratic elution utilizing a 7.8 x 300 mm Novapack C18 column (Waters), developed at 4 ml/min with 56% acetonitrile, 16% methanol, 28% 2% aqueous acetic acid (see Heyman et al., in *Cell* 68:1-20 (1992)). Absorbance was monitored between 200 and 600 nm using a Waters 996 photodiode array detector. Fractions were collected, dried and tested in the cotransfection assay for their ability to activate GAL-XOR6. Active fractions were pooled and rechromatographed on the same column using a gradient of methanol, 10mM ammonium acetate (pH 7.5) beginning at 30% methanol, run isocratically for 15 minutes, and then increasing linearly to 100% methanol over the next 45 minutes. Fractions were again tested for bioactivity and the active fractions pooled.

Final purification was accomplished using a dioxane/water gradient beginning at 20% dioxane and run isocratically for 15 minutes, then increasing linearly to 100% dioxane over the next 30 minutes.

Example 6
Ligand Binding

In order to investigate ligand binding, a protease protection assay was utilized (see Leng et al., 1993, supra, and Keidel et al, 1994, supra). ^{35}S-labelled protein was produced by coupled *in vitro* transcription/translation (TNT, Promega) and incubated with increasing concentrations of trypsin, chymotrypsin or alkaline protease in the presence of solvent carrier or with 10^{-5}M 3-amino ethylbenzoate (3-AEB) for 15 minutes at room temperature. The reactions were stopped with SDS-loading buffer and SDS-PAGE was performed on 12.5% acrylamide gels. Alterations in the size of protected fragments produced by added ligand in a dose dependent fashion was taken as evidence for specific binding.

3-AEB is seen to protect XOR-6 from trypsin digestion, thus confirming that 3-AEB binds XOR-6.

While the invention has been described in detail with reference to certain preferred embodiments thereof, it will be understood that modifications and variations are within the spirit and scope of that which is described and claimed.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Evans, Ronald M.
Blumberg, Bruce
Umesono, Kazuhiko

(ii) TITLE OF INVENTION: A NOVEL RXR-DEPENDENT SIGNALING PATHWAY AND LIGANDS USEFUL THEREFOR

(iii) NUMBER OF SEQUENCES: 3

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
(B) STREET: 444 South Flower Street, Suite 2000
(C) CITY: Los Angeles
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 90071

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/374,445
(B) FILING DATE: 17-JAN-1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Reiter, Stephen E.
(B) REGISTRATION NUMBER: 33,192
(C) REFERENCE/DOCKET NUMBER: P41 9887

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 619-546-4737
(B) TELEFAX: 619-546-9392

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2191 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 167..1324
(D) OTHER INFORMATION: /product= "XOR-6 RECEPTOR"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGAGAGTGAG AATCCCGGGC TCAGCCGCTC ACCTGTCCGG ATAGAGAGTT GGGATGTGAG 60

AGGGACAGAA GGGCGGGGCT AGTGCAGGTG TATCGGCCGC TCGAGGAGCT GCTCAGTGAA 120

```
AGAGAGAAGT GGCGAACGCT GGGACCAAGG TTTCTGTGAC AAAACG ATG TGG AAA          175
                                                    Met Trp Lys
                                                      1

GTG CAG GAG ACT TTG GTA CTG GAG GAA GAA GAG GAG GAA GAA GAC GCC          223
Val Gln Glu Thr Leu Val Leu Glu Glu Glu Glu Glu Glu Glu Asp Ala
  5                  10                  15

TCT AAC AGT TGT GGG ACG GGG GAA GAC GAG GAC GAT GGG GAC CCC AAG          271
Ser Asn Ser Cys Gly Thr Gly Glu Asp Glu Asp Asp Gly Asp Pro Lys
 20                  25                  30                  35

ATC TGC CGT GCG TGT GGG GAC CGG GCC ACT GGG TAT CAC TTC AAT GCT          319
Ile Cys Arg Ala Cys Gly Asp Arg Ala Thr Gly Tyr His Phe Asn Ala
                 40                  45                  50

ATG ACC TGC GAG GGC TGC AAG GGA TTC TTC AGG CGG GCC GTG AAG AGG          367
Met Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ala Val Lys Arg
            55                  60                  65

AAC TTG CGG CTC AGC TGC CCC TTC CAG AAT TCC TGC GTC ATC AAC AAG          415
Asn Leu Arg Leu Ser Cys Pro Phe Gln Asn Ser Cys Val Ile Asn Lys
        70                  75                  80

AGC AAT CGG CGC CAC TGC CAG GCC TGT CGG CTC AAG AAA TGT CTG GAC          463
Ser Asn Arg Arg His Cys Gln Ala Cys Arg Leu Lys Lys Cys Leu Asp
    85                  90                  95

ATC GGC ATG AGG AAA GAG TTG ATC ATG TCC GAT GCA GCG GTG GAA CAG          511
Ile Gly Met Arg Lys Glu Leu Ile Met Ser Asp Ala Ala Val Glu Gln
100                 105                 110                 115

AGA CGA GCG CTA ATT AAG AGA AAA CAC AAA TTA ACG AAA TTG CCC CCC          559
Arg Arg Ala Leu Ile Lys Arg Lys His Lys Leu Thr Lys Leu Pro Pro
                120                 125                 130

ACA CCC CCA GGG GCC AGT CTG ACT CCA GAG CAG CAG CAC TTT CTC ACT          607
Thr Pro Pro Gly Ala Ser Leu Thr Pro Glu Gln Gln His Phe Leu Thr
            135                 140                 145

CAA CTG GTT GGG GCC CAC ACC AAA ACC TTT GAC TTC AAC TTC ACC TTC          655
Gln Leu Val Gly Ala His Thr Lys Thr Phe Asp Phe Asn Phe Thr Phe
        150                 155                 160

TCC AAG AAC TTT CGG CCA ATA AGA AGA TCT TCA GAC CCA ACT CAG GAG          703
Ser Lys Asn Phe Arg Pro Ile Arg Arg Ser Ser Asp Pro Thr Gln Glu
    165                 170                 175

CCC CAA GCC ACC TCT TCT GAA GCC TTT TTG ATG CTA CCT CAT ATA TCT          751
Pro Gln Ala Thr Ser Ser Glu Ala Phe Leu Met Leu Pro His Ile Ser
180                 185                 190                 195

GAC CTC GTT ACC TAC ATG ATC AAG GGC ATC ATC AGC TTT GCC AAA ATG          799
Asp Leu Val Thr Tyr Met Ile Lys Gly Ile Ile Ser Phe Ala Lys Met
                200                 205                 210

CTC CCC TAC TTC AAG AGT CTG GAC ATT GAA GAC CAA ATT GCT CTC CTG          847
Leu Pro Tyr Phe Lys Ser Leu Asp Ile Glu Asp Gln Ile Ala Leu Leu
            215                 220                 225

AAA GGT TCT GTA GCG GAG GTT TCT GTG ATC CGA TTC AAC ACT GTG TTT          895
Lys Gly Ser Val Ala Glu Val Ser Val Ile Arg Phe Asn Thr Val Phe
        230                 235                 240

AAC TCT GAC ACC AAT ACG TGG GAG TGT GGC CCC TTC ACC TAT GAC ACT          943
Asn Ser Asp Thr Asn Thr Trp Glu Cys Gly Pro Phe Thr Tyr Asp Thr
    245                 250                 255
```

```
GAG GAT ATG TTC CTG GCC GGC TTC CGT CAG CTG TTC CTG GAG CCC CTG        991
Glu Asp Met Phe Leu Ala Gly Phe Arg Gln Leu Phe Leu Glu Pro Leu
260                 265                 270                 275

GTG AGG ATT CAT CGC ATG ATG AGG AAA CTG AAT GTA CAG AGT GAG GAA       1039
Val Arg Ile His Arg Met Met Arg Lys Leu Asn Val Gln Ser Glu Glu
                280                 285                 290

TAC GCC ATG ATG GCC GCT CTG TCC ATT TTC GCT TCT TAC CGA CCG GGG       1087
Tyr Ala Met Met Ala Ala Leu Ser Ile Phe Ala Ser Tyr Arg Pro Gly
            295                 300                 305

GTC TGC GAC TGG GAG AAG ATC CAG AAG CTG CAG GAA CAC ATT GCC CTG       1135
Val Cys Asp Trp Glu Lys Ile Gln Lys Leu Gln Glu His Ile Ala Leu
        310                 315                 320

ACA CTA AAA GAT TTC ATC GAC AGC CAA CGG CCC CCC TCC CCG CAG AAC       1183
Thr Leu Lys Asp Phe Ile Asp Ser Gln Arg Pro Pro Ser Pro Gln Asn
    325                 330                 335

AGG CTC CTG TAC CCC AAG ATC ATG GAG TGT CTG ACA GAG CTT CGG ACA       1231
Arg Leu Leu Tyr Pro Lys Ile Met Glu Cys Leu Thr Glu Leu Arg Thr
340                 345                 350                 355

GTC AAT GAC ATA CAC AGC AAG CAG CTC CTG GAG ATC TGG GAC ATT CAG       1279
Val Asn Asp Ile His Ser Lys Gln Leu Leu Glu Ile Trp Asp Ile Gln
                360                 365                 370

CCT GAT GCC ACC CCA CTT ATG CGA GAA GTC TTT GGA TCC CCT GAA           1324
Pro Asp Ala Thr Pro Leu Met Arg Glu Val Phe Gly Ser Pro Glu
            375                 380                 385

TGAGTGATGA GCACATTCCT ACTGTGAGAG TCGCTGACCC CACCGGGAAG CTTGGGCTCC     1384

TTCTACTGGC GTCTGTCCTG GTAGGGCAAT GTGGCCTTCA AAGCATCAGC AGCCGGTGGA     1444

TTGTCTTCTA CTGACACCAT CTTGTTCATT GCTCAGACGT TGCTTCAGTC CCATTGGGTC     1504

GAGGAGTTTA TGGAAAACTC TACCTTGTGG GATATCGGGG GGGGGAACAT GGAATTCCCA     1564

TCTGGGTCAC CAACATGTGA AAGAAACTGG TTCTGAGGAG CCAAAATGTT CTGCTGGACA     1624

AAAAGGAATG AAGTCACATA GAGACGAGTG TGGTCCAATA AAGAGACAGT CTGGCCAGAG     1684

ACAATGTGAC TGGTCCAATA TGAGTGGACA ATAAAGCAAC TCCCTGATCC TACAACTGGT     1744

TCCTGCAGGT TCTGCGCTGG GTTTGTGGCT CATTTAGATC AGGAGTTTGG TACCTGCACT     1804

AATTCTGTTC TTTTACGACT GACTCGGCTG AATGAAAGGG GCTGTCACTT GTAGCCGGCG     1864

ACGTGGGACA TTAGCCACAA GCCAAATCTT CTCAGGGAAG CCAAATGGGC TGGGGGGTGT     1924

AACACTGGGG GCACCAGACA AACTGTAACT AAATGAGGTT TAATCTCAGG GCTCCTGTAA     1984

TTATACTGAC CCCCCACTTG GGGATAGGGC TAAATATTGG GGGTCTGGGA GTTCTGTTCC     2044

AGAAGGTATT GGGGTGGGGG TCTATGGGTT GGGCCTGTGT TAGACGAGTG TTTGTAGCCG     2104

TTCCCTGTGT CTATTTAGTT CTGGTGTTTC TGGTACCGTA TTGGGCTCCA AATTGTTTTA     2164

TTCATAAAAA AAAAAAAAAA ACTCGAG                                         2191
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 386 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Trp Lys Val Gln Glu Thr Leu Val Leu Glu Glu Glu Glu Glu Glu
 1               5                  10                  15

Glu Asp Ala Ser Asn Ser Cys Gly Thr Gly Glu Asp Glu Asp Asp Gly
            20                  25                  30

Asp Pro Lys Ile Cys Arg Ala Cys Gly Asp Arg Ala Thr Gly Tyr His
        35                  40                  45

Phe Asn Ala Met Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ala
    50                  55                  60

Val Lys Arg Asn Leu Arg Leu Ser Cys Pro Phe Gln Asn Ser Cys Val
65                  70                  75                  80

Ile Asn Lys Ser Asn Arg Arg His Cys Gln Ala Cys Arg Leu Lys Lys
                85                  90                  95

Cys Leu Asp Ile Gly Met Arg Lys Glu Leu Ile Met Ser Asp Ala Ala
            100                 105                 110

Val Glu Gln Arg Arg Ala Leu Ile Lys Arg Lys His Lys Leu Thr Lys
        115                 120                 125

Leu Pro Pro Thr Pro Pro Gly Ala Ser Leu Thr Pro Glu Gln Gln His
    130                 135                 140

Phe Leu Thr Gln Leu Val Gly Ala His Thr Lys Thr Phe Asp Phe Asn
145                 150                 155                 160

Phe Thr Phe Ser Lys Asn Phe Arg Pro Ile Arg Arg Ser Ser Asp Pro
                165                 170                 175

Thr Gln Glu Pro Gln Ala Thr Ser Ser Glu Ala Phe Leu Met Leu Pro
            180                 185                 190

His Ile Ser Asp Leu Val Thr Tyr Met Ile Lys Gly Ile Ile Ser Phe
        195                 200                 205

Ala Lys Met Leu Pro Tyr Phe Lys Ser Leu Asp Ile Glu Asp Gln Ile
    210                 215                 220

Ala Leu Leu Lys Gly Ser Val Ala Glu Val Ser Val Ile Arg Phe Asn
225                 230                 235                 240

Thr Val Phe Asn Ser Asp Thr Asn Thr Trp Glu Cys Gly Pro Phe Thr
                245                 250                 255

Tyr Asp Thr Glu Asp Met Phe Leu Ala Gly Phe Arg Gln Leu Phe Leu
            260                 265                 270

Glu Pro Leu Val Arg Ile His Arg Met Met Arg Lys Leu Asn Val Gln
        275                 280                 285

Ser Glu Glu Tyr Ala Met Met Ala Ala Leu Ser Ile Phe Ala Ser Tyr
    290                 295                 300
```

```
Arg Pro Gly Val Cys Asp Trp Glu Lys Ile Gln Lys Leu Gln Glu His
305                 310                 315                 320

Ile Ala Leu Thr Leu Lys Asp Phe Ile Asp Ser Gln Arg Pro Pro Ser
                325                 330                 335

Pro Gln Asn Arg Leu Leu Tyr Pro Lys Ile Met Glu Cys Leu Thr Glu
            340                 345                 350

Leu Arg Thr Val Asn Asp Ile His Ser Lys Gln Leu Leu Glu Ile Trp
        355                 360                 365

Asp Ile Gln Pro Asp Ala Thr Pro Leu Met Arg Glu Val Phe Gly Ser
    370                 375                 380

Pro Glu
385
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGTTCATGAG AGTTCA 16

That which is claimed is:

1. A receptor polypeptide characterized by being responsive to the presence of hydroxy, mercapto or amino benzoate(s) to regulate the transcription of associated gene(s).

2. A polypeptide according to Claim 1 wherein said polypeptide is further characterized by having a DNA binding domain of about 66 amino acids with 9 Cys residues, wherein said DNA binding domain has about 73 % amino acid identity with the DNA binding domain of the human vitamin D receptor.

3. A polypeptide according to Claim 2 wherein said polypeptide is further characterized by having a ligand binding domain of about 203 amino acids, wherein said ligand binding domain has about 42 % amino acid identity with the ligand binding domain of the human vitamin D receptor.

4. A polypeptide according to Claim 1, wherein said polypeptide has substantially the same amino acid sequence as shown in SEQ ID NO:2.

5. A polypeptide according to Claim 1, wherein said polypeptide has the same amino acid sequence as shown in SEQ ID NO:2.

6. A heterodimer complex consisting of RXR and XOR-6.

7. Isolated DNA which encodes a polypeptide according to Claim 1.

8. DNA according to Claim 7 wherein said DNA encodes substantially the same amino acid sequence as shown in SEQ ID NO:2.

9. DNA according to Claim 7 wherein said DNA encodes the same amino acid sequence as shown in SEQ ID NO:2.

10. DNA according to Claim 7 comprising a segment having a contiguous nucleotide sequence which is substantially the same as nucleotides 166 - 1324 shown in SEQ ID NO:1.

11. DNA according to Claim 7 comprising a segment having a contiguous nucleotide sequence which is the same as nucleotides 166 - 1324 shown in SEQ ID NO:1.

12. A labeled single-stranded nucleic acid, comprising at least 20 contiguous bases in length having substantially the same sequence as any 20 or more contiguous bases selected from bases 1 - 2150, inclusive, of the DNA illustrated in SEQ ID NO:1, or the complement thereof.

13. A nucleic acid according to Claim 12 which is labelled with ^{32}P.

14. A nucleic acid according to claim 12 comprising at least 20 contiguous bases in length having substantially the same sequence as any 20 or more contiguous bases selected from bases 473 - 1324, inclusive, of the DNA illustrated in SEQ ID NO:1, or the complement thereof.

15. An isolated DNA construct comprising:
(i) the DNA of Claim 7 operatively linked to
(ii) regulatory element(s) operative for transcription of said DNA sequence and expression of said polypeptide in an animal cell in culture.

16. An animal cell in culture which is transformed with a DNA construct according to Claim 15.

17. A cell according to Claim 16, wherein said cell is further transformed with a reporter vector which comprises:
(a) a promoter that is operable in said cell,
(b) a hormone response element, and
(c) DNA encoding a reporter protein,
wherein said reporter protein-encoding DNA is operatively linked to said promoter for transcription of said DNA, and
wherein said promoter is operatively linked to said hormone response element for activation thereof.

18. An antibody which specifically binds a receptor polypeptide according to claim 1.

19. An antibody according to claim 18 wherein said antibody is a monoclonal antibody.

20. A method of making a receptor polypeptide according to claim 1, said method comprising culturing cells containing an expression vector operable in said cells to express a DNA sequence encoding said polypeptide.

21. A method according to Claim 20 wherein said receptor polypeptide has substantially the same amino acid sequence as shown in SEQ ID NO:2.

22. A method according to Claim 20 wherein said receptor polypeptide comprises a DNA binding domain with substantially the same sequence as that of amino acids 102 - 183 shown in SEQ ID NO:2.

23. A method according to Claim 20 wherein said DNA sequence comprises a segment with substantially the same nucleotide sequence as that of nucleotides 166 - 1324 shown in SEQ ID NO:1.

24. A method of identifying receptor polypeptide(s) characterized by being responsive to the presence of hydroxy, mercapto or amino benzoate(s) to regulate the transcription of associated gene(s), said method comprising hybridizing test DNA with a probe according to claim 14 under high stringency conditions, and selecting those sequences which hybridize to said probe.

25. A method of testing a compound for its ability to regulate transcription-activating effects of a receptor polypeptide according to claim 1, said method comprising assaying for the presence or absence of reporter protein upon contacting of cells containing said receptor polypeptide and reporter vector with said compound;

wherein said reporter vector comprises:

(a) a promoter that is operable in said cell,

(b) a hormone response element, and

(c) DNA encoding a reporter protein,

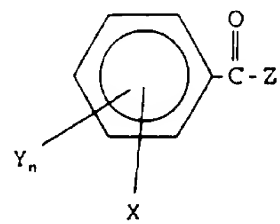
wherein said reporter protein-encoding DNA is operatively linked to said promoter for transcription of said DNA, and

wherein said promoter is operatively linked to said hormone response element for activation thereof.

26. A method according to Claim 25 wherein said contacting is carried out in the further presence of at least one hydroxy, mercapto or amino benzoate species.

27. A method for modulating process(es) mediated by receptor polypeptides according to claim 1, said method comprising conducting said process(es) in the presence of at least one hydroxy, mercapto or amino benzoate.

28. A method according to claim 27, wherein said amino benzoate is a compound having the structure:



wherein

X is a hydroxy, alkoxy, mercapto, thioalkyl, amino, alkylamino or acylamino group at the 2-, 3-, or 4-position of the ring,

each Y, when present, is independently selected from hydroxy, alkoxy, mercapto, thioalkyl, halide, trifluoromethyl, cyano, nitro, amino, carboxyl, carbamate, sulfonyl, sulfonamide,

Z is selected from -OR' or -NHR', wherein R' is selected from hydrogen, C_1-C_{12} alkyl or C_5-C_{10} aryl, and

n is 0-2.

29. A method according to claim 28 wherein X is 3-or 4-amino, Z is alkoxy and n is 0.

30. A method according to claim 29 wherein Z is selected from methoxy, ethoxy or butoxy.

31. A method according to claim 28 wherein X is 2-,3-, or 4-hydroxy, Z is alkoxy and n is 0.

32. A method according to claim 31 wherein Z is selected from methoxy, ethoxy or butoxy.

FIG. 1

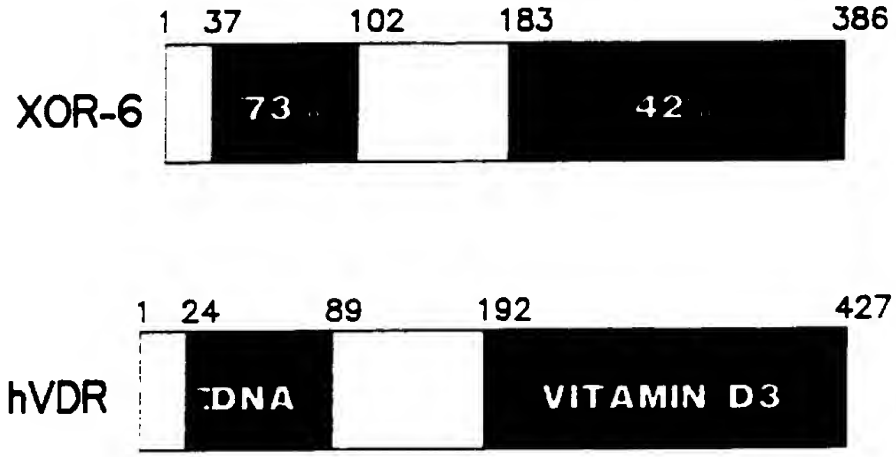
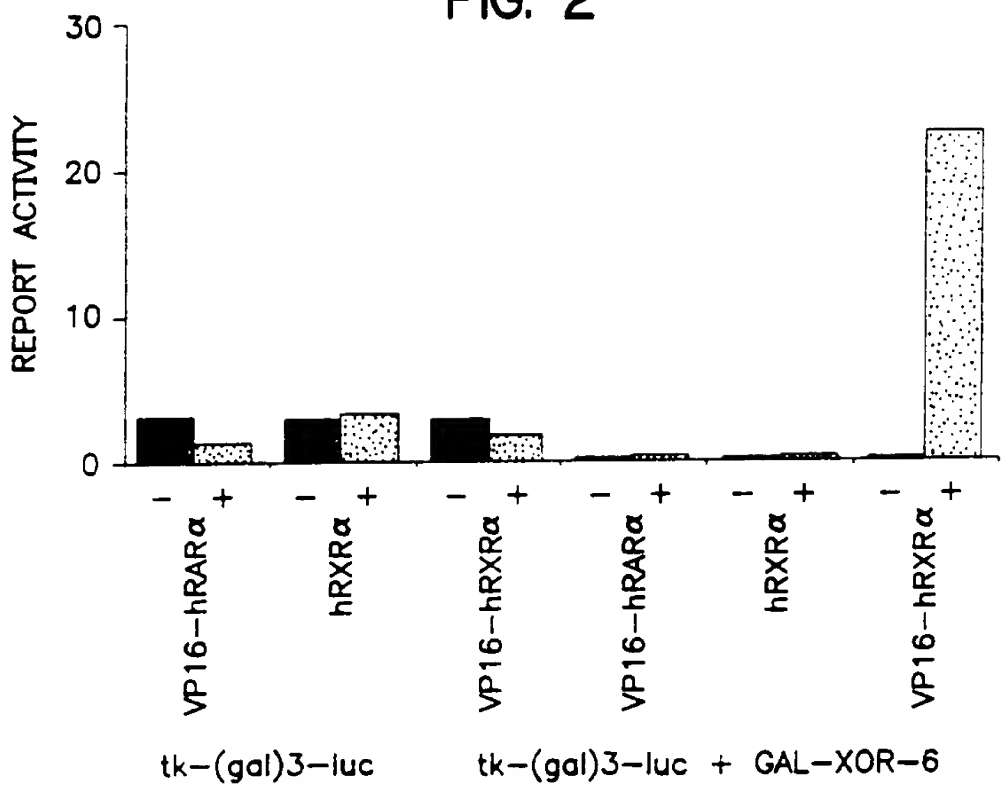

1
37
102
183
386
XOR-6
73
42
1
24
89
192
427
hVDR
DNA
VITAMIN D3

FIG. 2


30
20
10
0
REPORT ACTIVITY
– + – + – + – + – + – +
VP16–hRARα
hRXRα
VP16–hRXRα
VP16–hRARα
hRXRα
VP16–hRXRα
tk–(gal)3–luc
tk–(gal)3–luc + GAL–XOR–6

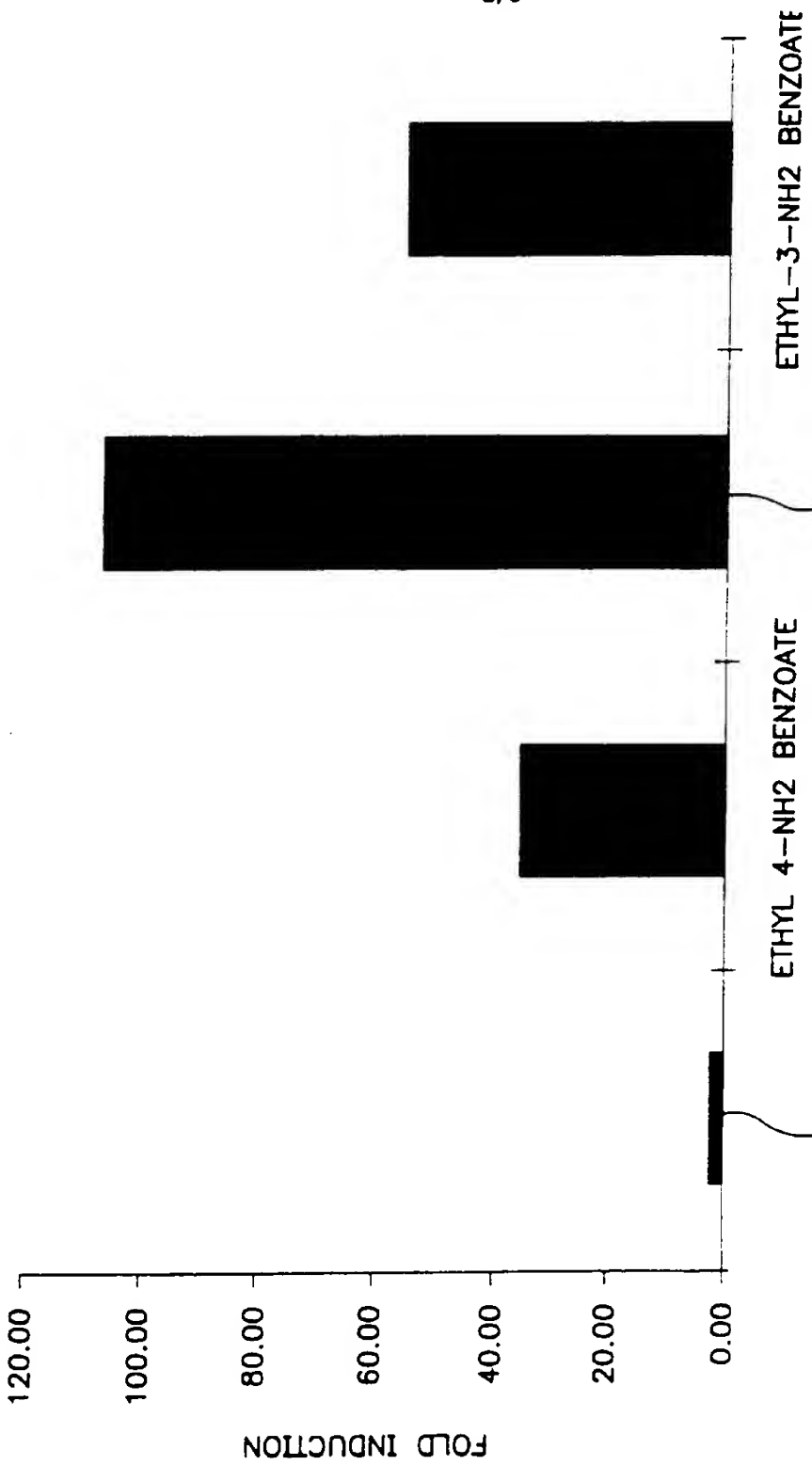

FOLD INDUCTION
120.00
100.00
80.00
60.00
40.00
20.00
0.00
METHYL 4-NH2 BENZOATE
ETHYL 4-NH2 BENZOATE
BUTYL 4-NH2 BENZOATE
ETHYL-3-NH2 BENZOATE
COMPOUND

FIG. 3

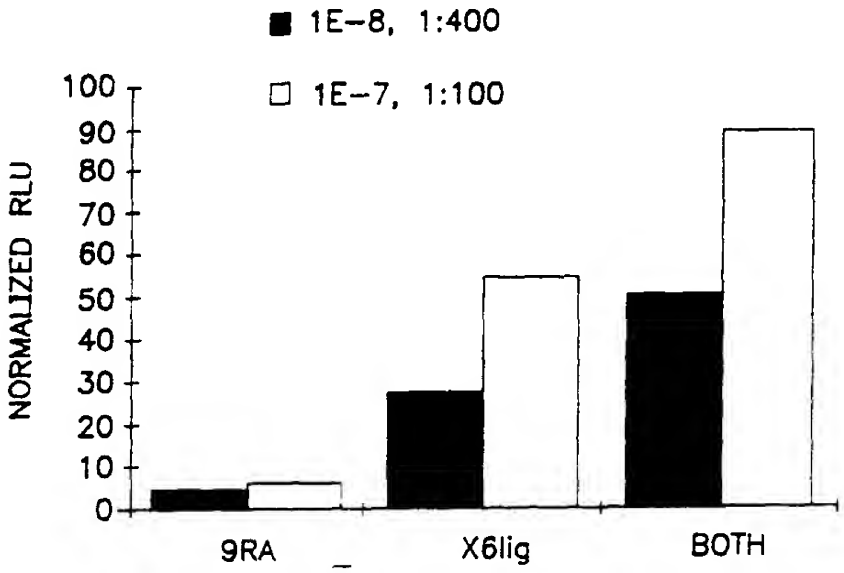

1E−8, 1:400
1E−7, 1:100
NORMALIZED RLU
100
90
80
70
60
50
40
30
20
10
0
9RA
X6lig
BOTH

FIG. 4

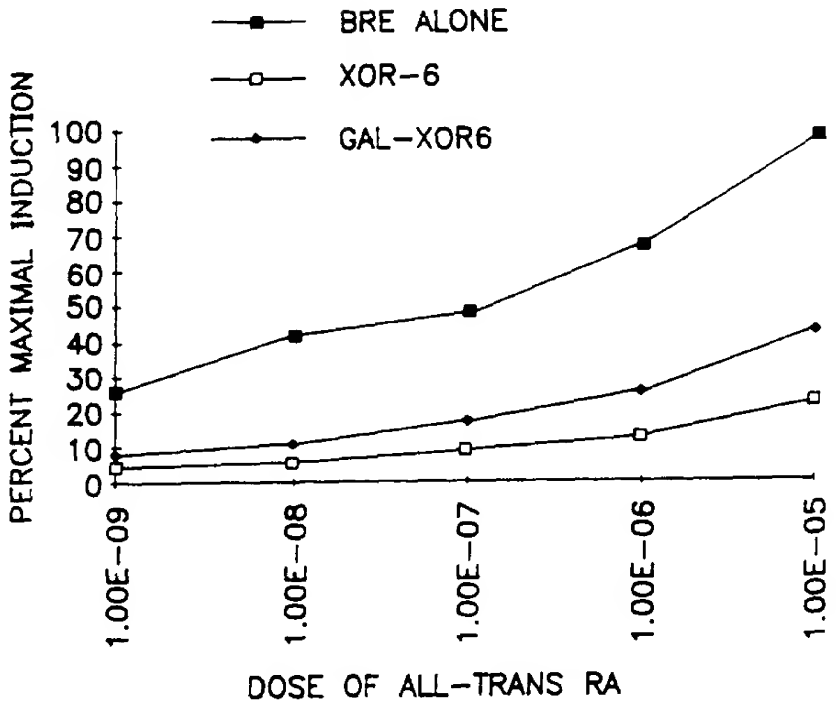

BRE ALONE
XOR−6
GAL−XOR6
PERCENT MAXIMAL INDUCTION
100
90
80
70
60
50
40
30
20
10
0
1.00E−09
1.00E−08
1.00E−07
1.00E−06
1.00E−05
DOSE OF ALL−TRANS RA

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/00058

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL : 435/6, 240.2; 530/350, 388.1; 536/23.1; 424/85.8
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 240.2; 530/350, 388.1; 536/23.1; 424/85.8

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAPLUS, MEDLINE, BIOSIS
search terms: Evans, R., Umesono, K., Blumberg, B., Xenopus, vitamin, steroid, receptor

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	SMITH et al. A Novel Receptor Superfamily Member in Xenopus that Associates with RXR, and Shares Extensive Sequence Similarity to the Mammalian Vitamin D3 Receptor. Nucleic Acids Research. 11 January 1994, vol. 22, pages 66-71, entire document	1-24 --- 25
Y	US 4,981,784 A (R.M. Evans) 01 January 1991, columns 14-20	25
A,P	MANGELSDORF et al. The RXR Heterodimers and Orphan Receptors. Cell .15 December 1995, vol 83, pages 841-850	1-32
A	EVANS, R. M. The Steroid and Thyroid Hormone Receptor Superfamily. Science. 13 May 1988, vol. 240, pages 889-895	1-32

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 APRIL 1996	24-05-96
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer AMY ATZEL Telephone No. (703) 308-0208

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/00058

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MCDONNELL, et al. Functional Domains of the Human Vitamin D3 Receptor Regulate Osteocalcin Gene Expression. Molecular Endocrinology. April 1989, vol 3, pages 635-644	1-32

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/00058

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C12Q 1/68; C12N 5/10; C07K 14/00, 2/00, 4/00, 14/00, 16/00; C07H 21/04; A61K 39/395